Chemistry & Biology, Vol. 11, 739–747, June, 2004, ©2004 Elsevier Ltd. All rights reserved. DOI 10.1016/j.chembiol.2004.06.005

Previews

An Early Taste of Functional Glycomics

Recent advances in the synthesis of carbohydrates and in microarrays for profiling biochemical activities have set the stage for rapid advances in understanding the biological roles for oligosaccharides. An early example by Seeberger and colleagues illustrates the use of these tools in HIV infection [1].

Following DNA and protein, carbohydrates are gaining recognition as a third class of biopolymer that plays an important role in a broad range of biological activities [2]. The development of DNA and protein science—and of the current emphasis on genomics and proteomics—can be traced to important advances in both the characterization and the synthesis of these biopolymers. Indeed, the development of rapid and reliable sequencing methods, together with synthetic protocols for the preparation of oligonucleotides and peptides on solid-phase, has made studies of protein and DNA biochemistry routine. The study of carbohydrates, by contrast, is only now benefiting from analogous developments.

On the synthesis front, two approaches are providing access to complex oligosaccharides. Peter Seeberger's group at the ETH has pioneered chemical methods for the automated synthesis of complex carbohydrates using a synthesizer analogous to those used for DNA and peptides [3]. Chi-Huey Wong's laboratory at the Scripps Research Institute has harnessed the enzymes that are involved in glycosylation to provide for biochemical synthesis of these molecules [4]. Taken together, these groups have provided routes to structures that were previously available only through many years of synthetic effort, and have laid the groundwork for further developments that will make the preparation of complex oligosaccharides routine.

On the analysis front, glycobiology faces a more difficult challenge than did protein and DNA research. First, oligosaccharides are often branched structures, complicating the task of structural characterization. Second, the carbohydrates are usually present as a heterogeneous population of related structures. Hence, it can be difficult to define the substructure of a complex carbohydrate that is the biologically relevant motif. And finally, there is no single enzymatic machinery for preparing carbohydrates, dimming the prospect for identification of a "glycomic" code that can aid the prediction of carbohydrate structures from genetic information.

Given these challenges, the practice of characterizing carbohydrates proceeds along two lines. Structural characterization follows the methods of chemical analysis, which begin with tedious purification of carbohydrate ligands and use of mass spectrometry, nuclear magnetic resonance spectroscopy, and other analytical methods to characterize the structures. The other interest is in characterizing the biological activities of the carbohydrates, which includes identifying receptors that bind the carbohydrates and enzymes that modify the structures.

The development of carbohydrate microarrays has the potential to vastly accelerate the biochemical characterization of carbohydrates. The microarrays are clearly inspired by the successes of the analogous gene chips and protein chips, which allow the identification of biochemical interactions from thousands of possibilities. The first carbohydrate arrays, reported in 2002, emphasized the preparation of the arrays and their utility in model experiments [5–7]. Since that time, several reports have disclosed approaches to carbohydrate arrays that differ in the source of the oligosaccharides (natural or synthetic), the chemistries used for immobilization of the structure, and the control over the densities, steric environment, and orientation of ligands. These reports have set the stage for studies that use the arrays to discover novel interactions in glyobiology.

The paper in this issue of *Chemistry & Biology* by Seeberger, O'Keefe, and coworkers makes an important advance by applying the microarrays to a biological problem of high interest: the gp120-mediated infection of HIV. In so doing, the work shows the value of the arrays in biomedical research and also points to opportunities for further developments in carbohydrate arrays.

The authors investigated the role of oligosaccharides in the infectious process of HIV. The viral surface presents a glycoprotein, gp120, which plays a prominent role in penetration of the virus into cells of the immune system. gp120-associated molecules interact with CD4 proteins and chemokine receptors on T lymphocytes, macrophages, and dendritic cells to initiate internalization of HIV by the host cell. These interactions depend on the N-linked high-mannose oligosaccharides present on the protein. Hence, an understanding of the role of the carbohydrates from gp120 in mediating infection could have important implications for the development of vaccines for treatment and the identification of protein targets for small molecule drug discovery.

Seeberger and colleagues used microarrays presenting oligosaccharides and glycoproteins to characterize the binding specificities of four gp120 binding proteins. These proteins included 2G12, a human monoclonal antibody, DC-SIGN, a dendritic cell lectin, and two proteins, cyanovirin and scytovirin from cyanobacteria that have been found to possess anti-HIV effects. Arrays presenting gp120, gp41 (another HIV-surface glycoprotein), and several other synthetic glycoproteins were used in experiments that applied fluorescently labeled proteins in order to identify protein-glycoprotein interactions. The aim of these studies was to identify the carbohydrate motifs that mediated the protein binding activities and to mechanistically distinguish the protein binding activities.

Among the results of these binding experiments was the finding that both DC-SIGN and 2G12 bind to the gp41 molecule. The results with 2G12 raise the question

of whether 2G12's in vivo efficacy is due to its interactions with gp120 alone or binding the high-mannose oligosaccharides present on both gp120 and gp41—an important question for designing synthetic vaccines. The authors also used the arrays to investigate the interactions of multiple proteins with gp120. This effort lead to the unexpected finding that pretreatment of gp120 with cyanovirin did not affect binding of CD4 to gp120, but when gp120 was first treated with CD4, addition of cyanovirin resulted in displacement of the CD4. Further, when CD4 was prebound to nonglycosylated gp120 (p120), cyanovirin did not displace CD4. The authors suggest a model wherein treatment of CD4-bound gp120 with cyanovirin induces conformational changes in gp120 that disrupt existing gp120-CD4 interactions. This intriguing finding may also be explained by different kinetic courses for binding that depend on the order of protein addition. In either event, the microarray-based experiments are significant because they identify unanticipated mechanisms for understanding CVN's inhibitory effects and in turn give rise to hypotheses that can be tested directly.

The authors next used a carbohydrate array to define the core motif that is recognized by the antibody. This information is vital to designing vaccines that give immune responses that mimic the 2G12 antibody. The array comprised a high-mannose nona-saccharide and several smaller oligosaccharides that are substructures of this parent. An analysis of the binding profile of 2G12 to this array revealed that the antigenic fragment comprised the disaccharide motif Manα1-2Man. Further studies established that the binding of antibody did not depend on the polypeptide backbone structure. The manuscript also describes analogous experiments to understand the binding profiles of other proteins, including scytovirin, a novel HIV-inactivating protein. Taken together, these examples establish that carbohydrate arrays can be used to rapidly profile the binding specificities of proteins, and should motivate many other researchers to adopt these tools.

Looking ahead, I believe that advances in carbohydrate microarray technology will impact two experimental objectives. The first concerns the identification of new interactions, particularly the enzymes that modify carbohydrates and the proteins that bind these structures. Current approaches with microarrays often start with a known enzyme or binding protein and apply the candidate to an array to identify the carbohydrate partners with which it interacts. These assays require labels—for example, a fluorescent group on the binding protein or a radiolabel that is enzymatically incorporated into the substrate—to identify the interactions. Hence, the strategies are best suited to identifying activities that are anticipated. More difficult, by contrast, is the identification of activities that are wholly unanticipated, and which therefore cannot be labeled. An exciting development that may enable these applications is the use of mass spectrometry methods to directly analyze carbohydrates and proteins on an array. One method uses self-assembled monolayers that present carbohydrates and has been shown to identify enzyme and protein binding activities [9].

A second objective is to carry out biochemical analysis of the interactions of carbohydrates with proteins and enzymes directly on a chip. The benefits of a chip-based format include the use of small amounts of reagents and rapid structure-activity profiling of the interaction for a family of carbohydrate analogs. These biochemical assays require that the microarrays provide quantitative information on activities, and that the data agree with the results of conventional solution-phase assays. There are now surface chemistries that provide good control over the orientations, densities, and environments of immobilized ligands and that also prevent nonspecific interactions with proteins. For carbohydrates, the control over density of ligand is especially important, since many proteins that bind these ligands do so in a polyvalent manner and can show dramatic differences in binding affinities and even specificities as the density is varied [10, 11].

The field of glycobiology has benefited greatly from advances in chemical tools—in both synthesis and analysis—over the past five years. With the paper by Seeberger and colleagues in this issue, these tools have been combined to better understand the role of the gp120 glycoprotein in HIV infection. We should expect that these approaches will be applied to several other glycobiology problems over the next period.

Milan Mrksich
The University of Chicago
Department of Chemistry
5735 South Ellis Avenue
Chicago, Illinois 60637

Selected Reading

1. Adams, E.W., Ratner, D.M., Bokesch, H.R., McMahon, J.B., O'Keefe, B.R., and Seeberger, P.H. (2004). Chem. Biol. *11*, this issue, 875–881.
2. Bertozzi, C.R., and Kiessling, L.L. (2001). Science *291*, 2357–2364.
3. Plante, O.J., Palmacci, E.R., and Seeberger, P.H. (2001). Science *291*, 1523–1527.
4. Fazio, F., Bryan, M.C., Blixt, O., Paulson, J.C., and Wong, C.-H. (2002). J. Am. Chem. Soc. *124*, 14397–14402.
5. Wang, D., Liu, S., Trummer, B.J., Deng, C., and Wang, A. (2002). Nat. Biotechnol. *20*, 275–281.
6. Houseman, B.T., and Mrksich, M. (2002). Chem. Biol. *9*, 443–454.
7. Fukui, S., Galustian, C., Lawson, A.M., and Chai, W. (2002). Nat. Biotechnol. *20*, 1011–1017.
8. Ratner, D.M., Adams, E.W., Su, J., O'Keefe, B.R., Mrksich, M., and Seeberger, P.H. (2004). Chembiochem *5*, 379–383.
9. Su, J., and Mrksich, M. (2002). Angew. Chem. Int. Ed. Engl. *41*, 4715–4718.
10. Shenoy, S.R., Barrientos, L.G., Ratner, D.M., O'Keefe, B.R., Seeberger, P.H., Groneneborn, A.M., and Boyd, M.R. (2002). Chem. Biol. *9*, 1109–1118.
11. Horan, N., Yan, L., Isobe, H., Whitesides, G., and Kahne, D. (1999). Proc. Natl. Acad. Sci. USA *96*, 11782–11786.